Toward robust visual odometry using prior 2D map information and multiple hypothesis particle filtering*

S. Edwards[1], L. Mihaylova[1], J. M. Aitken[1], and S. Anderson[1]

Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield, S1 3JD, UK.

Abstract. Visual odometry can be used to estimate the pose of a robot from current and recent video frames. A problem with these methods is that they drift over time due to the accumulation of estimation errors at each time-step. In this short paper we propose and briefly demonstrate the potential benefit of using prior 2D, top-down map information combined with multiple hypothesis particle filtering to correct visual odometry estimates. The results demonstrate a substantial improvement in robustness and accuracy over the sole use of visual odometry.

Keywords: Visual odometry · Deep learning · Multiple Hypothesis · Particle filter · Map prior.

1 Introduction

Visual odometry (VO) is a popular method of pose estimation in mobile robots and there are many methods for this including key-frame optimisation [8] and recently deep learning [6, 3]. The deep learning methods are advantageous because they avoid the need for camera calibration and online optimisation used in key-frame methods, although they tend to be less accurate than the optimisation methods.

One problem that is common to all VO methods (and indeed all odometry methods) is that the pose estimate drifts over time due to the accumulation of estimation errors. Yet, there is often additional information we can use to help reduce drift, such as prior map information. This is true in scenarios of driver-less cars (road maps), mobile robots moving along corridors in indoor environments (architectural floor-plans), and pipe inspection robots such as in the oil, gas, sewer/water and nuclear industries (where pipe network plans tend to be readily available). In all these cases we potentially have prior information in the form of a top-down 2D view of the map, and the mobile agent is largely constrained to move along routes in this map. So, it appears attractive to make

* This work is supported by the UK's Engineering and Physical Sciences Research Council (EPSRC) Programme Grant EP/S016813/1 Pervasive Sensing for Buried Pipes (Pipebots) `https://pipebots.ac.uk`

use of this information in scenarios where it is available. This idea is used in a VO system in [1], where the probability of being located in a discrete road segment is estimated using a particle filter.

The algorithm developed in this paper fuses VO with prior 2D map information using multiple hypothesis particle filtering: when a moving agent reaches a junction in the map, multiple particle filters are used to fuse the VO data with each possible route away from the junction. The most likely hypothesis is probabilistically selected using the distribution of particles for each filter, which acts as a likelihood function, similar to methods that have been used in multiple model particle filtering for fault detection [4]. The results demonstrate that simply using knowledge of the map alone with VO does not lead to successful pose estimation but instead a multiple hypothesis method must be used to ensure accuracy and robustness.

2 Methods

The proposed method of map hypothesis switching is intended for robots or vehicles operating in environments that highly constrain the agent's motion but whose exact layout is uncertain. Examples of such situations are road or pipe networks where the location of junctions and the connections between junctions are approximately known and available in the form of a 2D, top-down map.

The system performs VO via a deep network similar to those presented in [6, 5] using optic flow calculated via the Horn-Shunck method. The VO outputs are filtered via a particle filter that uses accelerometer and gyroscope data as the basis for a state-space model.

The map provided to the system is represented as a set of coordinates of known junctions with connections between each junction. The system uses the straight lines between each pair of connected junctions as its prior map. Multiple models are instantiated upon reaching a junction and we make one important assumption that the system can identify when it has reached a junction via a separate system. A standard method of multiple model particle filtering, as in [4], is adapted here to switch between particle filter models that express different hypotheses, where the distribution of particles for each filter acts as a likelihood function, and this is used to probabilistically select the most likely hypothesis.

We use the KITTI data set [2] consisting of camera and GPS data from a car, to both train the visual odometry and test the multi hypothesis system. We use different sequences for training and testing, to ensure testing is independent.

3 Results

In order to evaluate our proposed algorithm we tested 1. a system that used VO only, 2. a VO system using a map with a single particle filter, and 3. a VO system using a map with multiple hypothesis particle filtering on separate test data. Fig 1 shows the ground truth and pose estimates for each method overlaid on a street map. As can be seen, the raw VO system performs poorly and a single

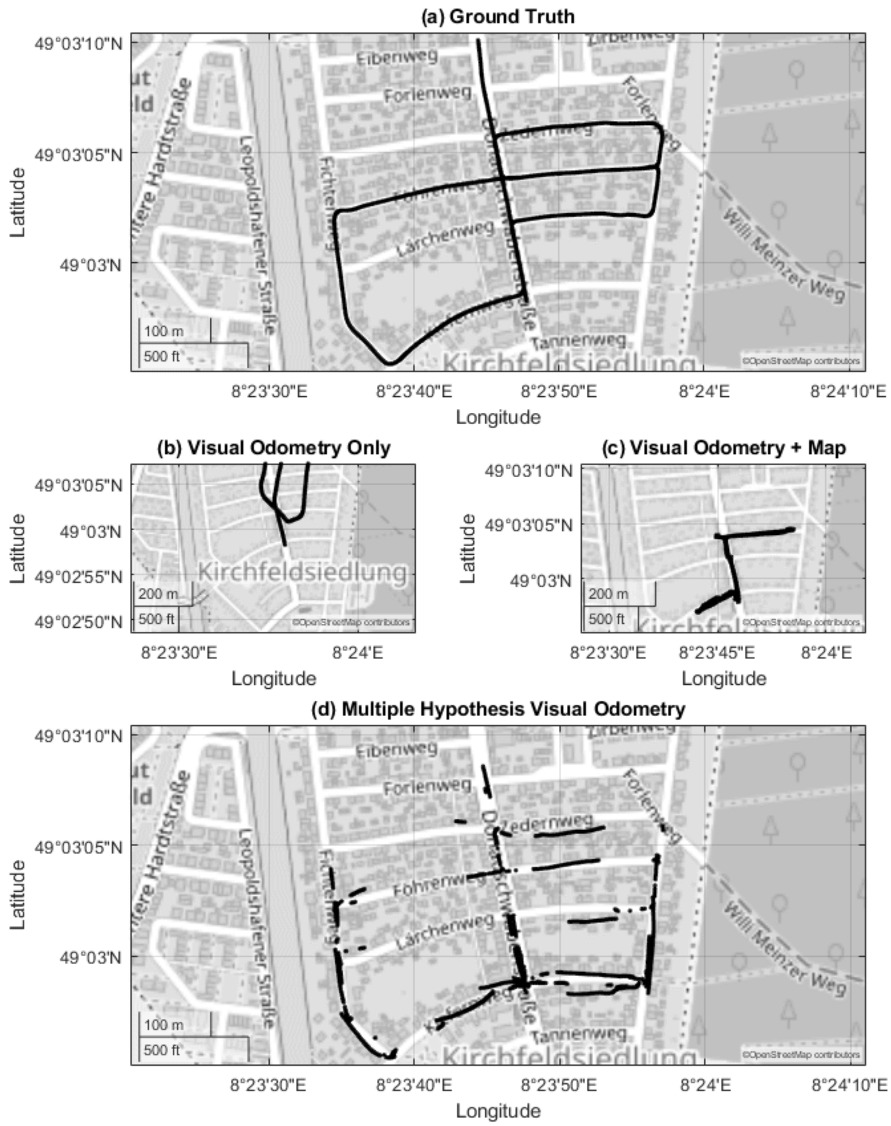

Fig. 1. Comparison of ground truth GPS vehicle data with pose estimation algorithms on independent validation data (KITTI, sequence 5). (a) Ground truth. (b) Visual odometry only (note that the pose estimate exits the area of the ground truth but we retain a comparable zoom-level to the other plots for clarity). (c) Visual odometry with map information. (d) Multiple hypothesis visual odometry.

map hypothesis results in unrecoverable failure when the system believes itself to be in the wrong section of the map. The multiple hypothesis system is generally much more accurate. This system occasionally selects an incorrect hypothesis (corresponding to the apparent gaps in location estimates in Fig. 1d) but then automatically recovers. The principal improvement of this system over simple odometry is the ability to recover from failure, however it also demonstrates an improved pose accuracy, with VO alone resulting in a mean position and heading error of 335.2m(343.20 and -1.65(-2.20) radians respectively and a final positional error of 230.4 m, and the multi hypothesis system resulting in mean errors of 49.2m (40.3) and 0.28 (0) radians and a final error of 0.85 m.

4 Conclusions

In this work we have presented a novel VO method that uses prior 2D map information and multiple hypothesis particle filtering. We demonstrated that the method was more accurate and robust than solely using VO, and that using multiple hypothesis particle filtering substantially improved on using a single particle filter with map information, particularly in it's ability to recover from errors. In future work we aim to resolve the problem of junction recognition in order to make a fully standalone algorithm and also develop a compact system that can run in real-time on mobile hardware.

Additionally we aim to address the system's main weakness, that of temporarily selecting an incorrect hypothesis, which occurs regularly after junctions. Improvements here may come from including additional information in the probability calculations, such as designed or learned features from each hypotheses parameters similar to [7], or from more complex assessments of each hypotheses probabilities compared to each other and possibly the system's states and sensor inputs.

References

1. Brubaker, M.A., Geiger, A., Urtasun, R.: Map-based probabilistic visual self-localization. IEEE Transactions on Pattern Analysis and Machine Intelligence **38**(4), 652–665 (2016)
2. Geiger, A., Lenz, P., Stiller, C., Urtasun, R.: Vision meets robotics: The KITTI dataset. The International Journal of Robotics Research **32**(11), 1231–1237 (2013)
3. Han, L., Lin, Y., Du, G., Lian, S.: DeepVIO: Self-supervised Deep Learning of Monocular Visual Inertial Odometry using 3D Geometric Constraints. In: 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 6906–6913 (2019)
4. Kadirkamanathan, V., Li, P., Kirubarajan, T.: Sequential Monte Carlo filtering vs. the IMM estimator for fault detection and isolation in nonlinear systems. In: Willett, P.K., Kirubarajan, T. (eds.) Component and Systems Diagnostics, Prognosis, and Health Management. vol. 4389, pp. 263 – 274. International Society for Optics and Photonics, SPIE (2001)

5. Konda, K., Memisevic, R.: Learning visual odometry with a convolutional network. In: The 10th International Conference on Computer Vision Theory and Applications (VISAPP). pp. 486–490 (2015)
6. Muller, P., Savakis, A.: Flowdometry: An optical flow and deep learning based approach to visual odometry. In: 2017 IEEE Winter Conference on Applications of Computer Vision (WACV). pp. 624–631. IEEE (2017)
7. Ossig, D.L., Kurzenberger, K., Speidel, S.A., Henning, K.U., Sawodny, O.: Sensor fault detection using an extended kalman filter and machine learning for a vehicle dynamics controller. In: IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society. pp. 361–366. IEEE (2020)
8. Scaramuzza, D., Fraundorfer, F.: Visual Odometry [Tutorial]. IEEE Robotics & Automation Magazine **18**(4), 80–92 (2011)